

CLAIMS

We claim:

1. A method of computing response time of a web server comprising the steps of:
placing a plurality of correlation tags in data at networking and application layers,
wherein said tags allow for later identification of said data;
collecting said data from said layers, wherein said data corresponds to a single event;
combining said data from networking and application layers into a metric, wherein said data corresponding to a single web event is identified; and
calculating client perceived response time.
2. The method of claim 1, wherein said event is selected from one of a web page download and a web session.
3. The method of claim 1, wherein said calculating step is performed using analytical models of response time.
4. A method of estimating of a perceived response time of at least one web server computing device to one or more client computing devices connected to the at least one web server device via a network, the method comprising the steps of:
generating and placing a session identifier (ID) as a correlation tag in each of a plurality of requests sent by a client to a web server device, wherein said correlation tags identify said requests;
generating and placing a connection identifier (ID) as a correlation tag in each communication packet sent between the client and the web server device;
combining said plurality of requests and said communication packets into a metric, wherein said each request and communication packet corresponding to a single event is identified; and
estimating client perceived response time of at least one web server computing device to a request by one or more client computing devices connected to the web server device via a network.

5. The method of claim 4, wherein the network is the Internet.

6. The method of claim 4, wherein said step of generating and placing the session ID further comprises a step of establishing a web session between the client and the web server device.

7. The method of claim 4, further comprising a step of logging each web session between the client and the web server device.

8. The method of claim 4, wherein said step of generating and placing a connection ID further comprises a step of establishing a network connection between the client and the web server device.

9. The method of claim 4, further comprising a step of logging said each communication packet sent between the client and the web server.

10. The method of claim 4, further comprising a step of grouping all of said plurality of requests and said communication packets corresponding to a single event.

11. The method of claim 4, wherein said estimating step further comprises the steps of

- a) retrieving a page composition vector and TCP/IP round trip time (RTT), packet loss rate, and average connection time T_c ;
- b) calculating time $T1 = T_c + C1(b1)$ and time $T2 = C1(o) + T_c + C2(b2)$ and setting a loop counter;
- c) averaging $T1$ and $T2$ by $(T1 + T2)/2$ and terminating processing if the loop counter is less than or equal to the value n ;
- otherwise, if $T1$ is smaller than $T2$, $T1$ is set to $T1 + (RTT/2) + C1(BI)$, and
- if $T1$ is not smaller than $T2$, $T2$ is set to $T2 + (RTT/2) + C2(BI)$; and
- d) incrementing the loop counter and repeating step c,

wherein the page composition vector is composed of $\{b1, b2, \dots, bn\}$ and o ,

b_i , where i is an number 1,2,..., being the size of the i -th component of the web page,

n being the number of components,

o being the offset at which first component is embedded in a container page,

C1(y) being the time it takes to download y bytes on a first TCP/IP connection between the client and the web server device, and

C2(y) being the time it takes to download y bytes on a second TCP/IP connection between the client and web server device.

12. An apparatus for estimating a perceived client response time of at least one web server computing device to one or more client computing devices connected to the at least one web server via a network, the method comprising:

a means for generating and placing a session identifier (ID) as a correlation tag in each of a plurality of requests sent by the one or more client computing devices to the at least one web server, wherein said correlation tags identify said requests;

a means for generating and placing a connection identifier (ID) as a correlation tag in each communication packet sent between the one or more client computing devices and the at least one web server;

a means for combining said plurality of requests and said communication packets into a metric, wherein said each request and communication packet corresponding to a single event is identified; and

a means for estimating the perceived client response time of at least one web server computing device to a request by one or more client computing devices connected to the web server via a network.

13. A computer program device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform method steps for an apparatus for estimating a perceived client response time of at least one web server computing device to one or more client computing devices connected to the at least one web server via a network, the method comprising:

a means for generating and placing a session identifier (ID) as a correlation tag in each of a plurality of requests sent by the one or more client computing devices to the at least one web server, wherein said correlation tags identify said requests;

a means for generating and placing a connection identifier (ID) as a correlation tag in each communication packet sent between the one or more client computing devices and the at least one web server;

a means for combining said plurality of requests and said communication packets into a metric, wherein said each request and communication packet corresponding to a single event is identified; and

a means for estimating the perceived client response time of at least one web server computing device to a request by one or more client computing devices connected to the web server via a network.